2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Graphon Field Tracking Games with Discrete Time Q-noise

Alex Dunyak, *Member, IEEE*, and Peter E. Caines, *Life Fellow, IEEE*

***Abstract*— Linear quadratic games on very large dense networks can be modelled with discrete time linear quadratic graphon field games with Q-noise. In such a game, each node in the graph corresponds to an agent weakly connected via an undirected network, with a correlated Brownian disturbance affecting each agent. The limit of the finite-sized linear quadratic network tracking game in discrete time is formulated, and it is shown that under the proper assumptions, the game has a graphon limit system with Q-noise. Then, the optimal control of the discrete time system is found in closed-form and the Nash equilibrium behavior of the game is numerically demonstrated.**

I. Introduction

Large systems composed of interacting non-cooperative agents arise in many applications such as cellular networks, financial markets, and electrical networks. Modelling and control of such systems can be difficult or intractable due to their size and the complexity of their respective networks.

When the agents are uniform and their network completely connected one can use Mean Field Game theory to find the approximate system behavior by simulating a system with an infinite number of agents and finding the distribution of agents' states under different control methods ([1], [2], [3]). Standard mean field games can be extended to games on networks where each node has an infinite population of players by the use of graphons ([4], [5], [6]). Graphons [7] are a limit object for graphs, allowing the adjacency matrix of an infinitely large graph to be represented as an $L_2([0,1])$ integral operator. Under the Graphon Mean Field Games (GMFG) model each node in a network contains a separate, infinite population interacting with the agents local to their node uniformly and the agents in the network through the graphon. As there are an infinite number of agents, the actions of a single agent do not change the mean field.

In this work, a linear quadratic game on an infinitely large dense graph is investigated where each node represents a single agent, and where there is a correlated Gaussian noise affecting the agents. A continuous time, deterministic model with stochastic initial conditions for this type of linear-quadratic game was investigated by Gao, Foguen-Tchuendom, and Caines [8]. To distinguish this from the infinite-agent-per-node GMFG model, this approach was termed the Graphon Field Game model. As in the GMFG model, the actions of any individual agent do not directly affect the field of the system. The Nash equilibria of such a system requires the consistency of each agent's chosen actions with the field generated by each agent's optimal strategy, rather than the verification that no agent benefits by deviating from their current strategy.

This work extends the work of Gao, Foguen-Tchuendom, and Caines [8] by applying the Q-noise foundations of Dunyak and Caines [9] to discrete time systems. The expansion of results to discrete time allows the Nash Equilibrium of the game to be precisely simulated numerically. This model is analogous to the limit behavior of a finite dimensional graph system with a correlated Gaussian disturbance impacting each node at each time step. It is demonstrated that the discrete time linear quadratic Q-noise tracking game has an adapted Nash equilibrium solution, and the behavior of the equilibrium solution is shown numerically.

II. Preliminaries

A. Notation

For any function $f(\cdot)$, $f^*(\cdot)$ denotes the adjoint of f.

Consider the set of bounded symmetric non-negative functions $\mathbf{Q} : [0,1]^2 \to \mathbb{R}$. The function $\mathbf{Q}$ serves as the covariance of a stochastic process. The class of such functions denoted $\mathcal{Q}$ consists of those for which the following inequality is satisfied for every function $\boldsymbol{f} \in \mathcal{L}^2[0,1]$,

$$0 \le \int_0^1 \int_0^1 \mathbf{Q}(x,y)\boldsymbol{f}^*(x)\boldsymbol{f}(y)dxdy < \infty. \tag{1}$$

For any $\boldsymbol{f}, \boldsymbol{g}$ in the Hilbert space $\mathcal{L}^2[0,1]$, the inner product is defined via the Lebesgue integral,

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \boldsymbol{f}^* \boldsymbol{g} = \int_0^1 \boldsymbol{f}(x)\boldsymbol{g}(x)dx. \tag{2}$$

The linear integral operator using $\mathbf{Q} \in \mathcal{Q}$ as a kernel acting on a function $\boldsymbol{f} \in \mathcal{L}^2[0,1]$ is defined by

$$(\mathbf{Q}f)(x) = \int_0^1 \mathbf{Q}(x,y)\boldsymbol{f}(y)dy, \quad \forall\ x \in [0,1]. \tag{3}$$

The identity operator is denoted $\mathbb{I}$.

This article focuses on systems on games where each agent possesses a scalar state. The extension to vectors of states is straightforward, though it may add some technical considerations with regard to the commutation of operators.

B. Discrete Time Q-noise Processes

Discrete time Q-noise processes are $L_2([0,1])$ valued random processes satisfying the following axioms (modified from [9] for discrete time):

The authors are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail:alex.dunyak@mail.mcgill.ca; peterc@cim.mcgill.ca).

Work supported by *A. Dunyak NSERC 2019-05336, **P.E. Caines: NSERC 2019-05336, ARL W911NF1910110, Air Force OSR Grant FA9550-23-1-0015

1) Let $\mathbf{Q} \in \mathcal{Q}$, and let $([0,1] \times \{0,1,...,T\} \times \Omega, \mathcal{B}([0,1] \times \{0,1,...,T\} \times \Omega), \mathbb{P})$ be a probability space with the measurable random variable $\boldsymbol{g}_k(\alpha, \omega) : [0,1] \times \{0,1,...,T\} \times \Omega \to \mathbb{R}$ for all $k \in \{0,1,...,T\}, \alpha \in [0,1]$, $\omega \in \Omega$. For notation, ω is suppressed when the meaning is clear.
2) For all $\alpha \in [0,1]$, $\boldsymbol{g}_k(\alpha) \sim \mathcal{N}(0, \mathbf{Q}(\alpha, \alpha))$.
3) For all α and β, $\mathbb{E}[\boldsymbol{g}_k(\alpha)\boldsymbol{g}_k(\beta)] = \mathbf{Q}(\alpha, \beta)$.

An orthonormal basis example: Let $\{W_k^1, W_k^2, \cdots\}$ be a sequence of independent standard normal random variables for each $k \in \{0,1,...,T\}$. Let $\mathbf{Q} \in \mathcal{Q}$ have a diagonalizing orthonormal basis $\{\phi_r\}_{r=1}^{\infty}$ with eigenvalues $\{\lambda_r\}_{r=1}^{\infty}$. Then

$$\boldsymbol{g}_k(\alpha, \omega) = \sum_{r=1}^{\infty} \sqrt{\lambda_r}\phi_r(\alpha) W_k^r(\omega) \tag{4}$$

is a discrete time Q-noise process. For general Gaussian measures in function spaces, see e.g. [10].

III. Problem Statement

A. Discrete-time network system games

Consider a discrete time system on a graph $G^N = (V^N, E^N)$ where each node i represents an agent. The state of agent i at time k (denoted x_k^i with control u_k^i) evolves with the following stochastic difference equation:

$$x_{k+1}^i = (Ax_k^i + Bu_k^i + \frac{1}{N}\sum_{j=1}^{N} M_{ij}^N x_k^j) + W_k^i \tag{5}$$

where $A, B \in \mathbb{R}$, M^N is the weighted adjacency matrix of G^N, and $\{W_k^i\}$ is a collection of Gaussian disturbances with covariance matrix Q^N for each k. Subject to the actions of all other agents, each agent i minimizes the expected quadratic cost function with respect to their information set $\mathcal{F}_k^i$,

$$J^i(u^i, u^{-i}|\{z_k^i\}_{i=1}^N) = \mathbb{E}\left[\sum_{k=0}^{T} ||x_k^i - z_k^i||_S^2 + ||u_k^i||_R^2 \middle| \mathcal{F}_k^i\right], \tag{6}$$

where $z_k^i = \frac{1}{N}\sum_{j=1}^{N} M_{ij}^N x_k^j$, $||v||_S^2 = v^* S v$ for some $S \in \mathbb{R}$, $R \in \mathbb{R}$, $S \geq 0$ and $R > 0$.

There exists a Nash Equilibrium of the game when no agent can benefit by deviating from their current strategy. If the optimal strategy tuple is $\{u^{i*}\}_{i=1}^N$, this implies

$$J^i(u^{i*}, u^{-i*}) \leq J^i(u^i, u^{-i*}) \quad \forall i \in \{1, ..., N\}. \tag{7}$$

As the network size grows, the networked system adjacency matrix M^N approaches its associated graphon which is a bounded measurable function mapping $[0,1] \times [0,1] \to [0,1]$, denoted $\boldsymbol{M}$ (see [11], [7]). When the underlying graph is undirected, its graphon is also symmetric. An example of this network convergence is shown with the two finite networks in Fig. 1 converging to the graphon limit (Fig. 2).

B. Graphon Field Tracking Games

By letting the size of the network approach infinity, each agent in the system is associated with a point α on the unit interval. Define the discrete time Q-noise $\boldsymbol{g}_k^\alpha$, $\alpha \in [0,1]$, and the resulting discrete time system evolves according to

$$\boldsymbol{x}_{k+1}^\alpha = (A\boldsymbol{x}_k^\alpha + B\boldsymbol{u}_k^\alpha + D\boldsymbol{z}_k^\alpha) + \boldsymbol{g}_k^\alpha, \tag{8}$$

$$\boldsymbol{z}_k^\alpha = \int_0^1 \boldsymbol{M}(\alpha, \beta)\boldsymbol{x}_k^\beta d\beta \quad \forall\, \alpha \in [0,1]. \tag{9}$$

The local field for an agent designated by α refers to the value $\boldsymbol{z}_k^\alpha$ found using the above integral.

The objective function for agent α has the limit

$$J^\alpha(\boldsymbol{u}^\alpha, \boldsymbol{x}_0) = \mathbb{E}\left[\sum_{k=0}^{T} ||\boldsymbol{x}_k^\alpha - \boldsymbol{z}_k^\alpha||_S^2 + ||\boldsymbol{u}_k^\alpha||_R^2 \middle| \mathcal{F}_k^\alpha\right]. \tag{10}$$

As with mean field games and graphon mean field games, the graphon field term $\boldsymbol{z}_k^\alpha$ is asymptotically independent of both the state $\boldsymbol{x}_k^\alpha$ and the action $\boldsymbol{u}_k^\alpha$ of any other single agent. As with many mean-field game problems, this changes the limit problem from a game to a tracking control problem where each node in the network is penalized for deviating from its associated graphon field.

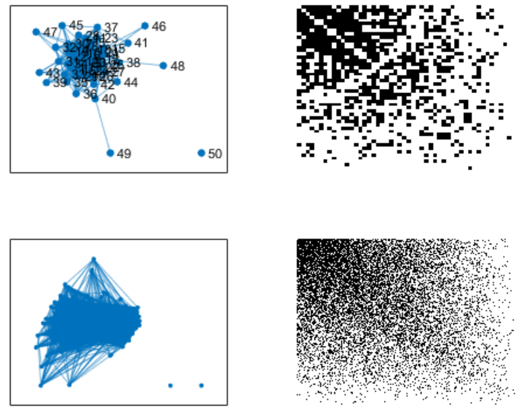

Fig. 1. Graphs of graphs with 50 and 500 nodes, respectively, where their associated adjacency matrices converge to the graphon in Fig. 2 when mapped to the unit square. Lower indexed nodes are more likely to be connected than higher indexed nodes.

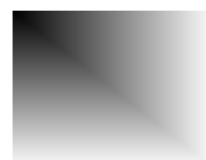

Fig. 2. The graph sequence shown in Fig. 1 converges to the graphon $\boldsymbol{W}(\alpha, \beta) = 1 - \max(\alpha, \beta), \quad \alpha, \beta \in [0,1]$.

IV. Solution to the Q-noise Graphon Field Game

The objective of agent α in the limit graphon field game is to minimize the following functional,

$$J^\alpha(\boldsymbol{u}^\alpha, \boldsymbol{x}_0) = \mathbb{E}\left[\sum_{k=1}^{T} ||\boldsymbol{x}_k^\alpha - \boldsymbol{z}_k^\alpha||_S^2 + ||\boldsymbol{u}_k^\alpha||_R^2 \middle| \mathcal{F}_k^\alpha\right], \tag{11}$$

where

$$\boldsymbol{x}_{k+1}^{\alpha} = (A\boldsymbol{x}_k^{\alpha} + B\boldsymbol{u}_k^{\alpha} + D\boldsymbol{z}_k^{\alpha}) + \boldsymbol{g}_k^{\alpha}, \tag{12}$$

and $\boldsymbol{x}_k \in L_2([0,1])$, $\boldsymbol{z}_k^{\alpha} = \int_0^1 \boldsymbol{M}(\alpha,\beta)\boldsymbol{x}_k^{\beta} d\beta$, $\boldsymbol{z}_k \in L_2([0,1])$, T is an integer representing the terminal time step, and u_k is assumed to be adapted to the information set $\mathcal{F}_k$.

For the simplicity of notation of the following sections, the filtration $\mathcal{F}_k$ will be omitted from the expectation, i.e. $\mathbb{E}_k[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_k^{\alpha}]$. The game is solved in two steps, first by formulating the response of an individual agent $\alpha \in [0,1]$ as a stochastic tracking problem, then by showing that the individual actions of each agent generate a Nash equilibrium.

A. Solution to the Stochastic Control Tracking Problem

The first step is to solve the stochastic tracking problem, where a given agent α is tracking an exogenous square-integrable drift signal $\boldsymbol{v}_k(\alpha),\ \alpha \in [0,1]$. The value function is found using dynamic programming,

$$\begin{aligned} V_k^{\alpha}(\mathcal{F}_k^{\alpha}) = \mathbb{E}\big[&||\boldsymbol{x}_k^{\alpha} - \boldsymbol{v}_k^{\alpha}||_S^2 + ||\boldsymbol{u}_k^{\alpha}||_R^2 \\ &+ V_{k+1}(\mathcal{F}_{k+1}^{\alpha})|\mathcal{F}_k^{\alpha}\big], \end{aligned} \tag{13}$$

$$V_M^{\alpha}(\mathcal{F}_M^{\alpha}) = ||\boldsymbol{x}_k^{\alpha} - \boldsymbol{v}_k^{\alpha}||_S^2. \tag{14}$$

This work considers the case where all agents have the full-information set $\mathcal{F}_k$ consisting of $\boldsymbol{x}_k^{\eta}$ and $\boldsymbol{v}_k^{\eta}$ for all $\eta \in [0,1]$. For simplicity, the scalar case is analyzed, where $\boldsymbol{x}_k^{\alpha}$, $\boldsymbol{u}_k^{\alpha}$, and $\boldsymbol{v}_k^{\alpha}$ are real scalar values for all $\alpha \in [0,1]$.

Lemma 4.1: The value function of agent α at time k, V_k^{α}, is given by

$$\begin{aligned} V_k^{\alpha}(\boldsymbol{x}_k) =&\mathbb{E}_k[(\boldsymbol{x}_k^{\alpha})^* P_k(\boldsymbol{x}_k^{\alpha}) + 2(\boldsymbol{x}_k^{\alpha})^*\boldsymbol{s}_k^{\alpha} + m_k^{\alpha}], \\ &k = \{0,...,T\}, \end{aligned} \tag{15}$$

where P_k is an positive scalar, and $\boldsymbol{s}_k$ and m_k^{α} are $L_2([0,1])$ valued functions for all $k = \{0,...,T\}$ derived from the following backwards recurrence relations,

$$F_k = (R + B^* P_{k+1} B)^{-1} B^* P_{k+1} A, \tag{16}$$

$$G_k = (R + B^* P_{k+1} B)^{-1} B^* P_{k+1} D, \tag{17}$$

$$H_k = (R + B^* P_{k+1} B)^{-1} B^*, \tag{18}$$

$$P_k = S + F_k^* R F_k + (A - BF_k)^* P_{k+1}(A - BF_k), \tag{19}$$

$$\begin{aligned} \boldsymbol{s}_k^{\alpha} =& - S\boldsymbol{v}_k^{\alpha} + F_k^* R(G_k\boldsymbol{v}_k^{\alpha} + H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]) \\ &+ \frac{1}{2}(A - BF_k)^* P_{k+1} \\ &\cdot \left[(D - BH_k)\boldsymbol{v}_k^{\alpha} - BH_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]\right] \\ &+ (A - BF_k)^*\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}], \end{aligned} \tag{20}$$

$$\begin{aligned} m_k^{\alpha} =& \boldsymbol{v}_k^{\alpha *} S\boldsymbol{v}_k^{\alpha} + (G_k\boldsymbol{v}_k^{\alpha} + H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}])^* R \\ &\cdot (G_k\boldsymbol{v}_k^{\alpha} + H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]) \\ &+ \left[(D - G_k)\boldsymbol{v}_k^{\alpha} - BH_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]\right]^* P_{k+1} \\ &\cdot \left[(D - G_k)\boldsymbol{v}_k^{\alpha} - BH_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]\right] \\ &+ 2\left[(D - G_k)\boldsymbol{v}_k^{\alpha} - BH_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]\right]\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}] \\ &+ \mathbf{Q}(\alpha,\alpha) + \mathbb{E}_k[m_{k+1}^{\alpha}], \end{aligned} \tag{21}$$

with the terminal conditions

$$P_T = S, \tag{22}$$

$$\boldsymbol{s}_T^{\alpha} = - S \cdot \boldsymbol{v}_T^{\alpha}, \tag{23}$$

$$m_T^{\alpha} = S \cdot |\boldsymbol{v}_T^{\alpha}|^2. \tag{24}$$

Further, the optimal control is given by

$$\begin{aligned} u_k^{o,\alpha} =& - (R + B^* P_{k+1} B)^{-1} B^* [P_{k+1}(A\boldsymbol{x}_k^{\alpha} + D\boldsymbol{v}_k^{\alpha}) \\ &+ \mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]] \end{aligned} \tag{25}$$

$$=: - F_k\boldsymbol{x}_k^{\alpha} - G_k\boldsymbol{v}_k^{\alpha} - H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]. \tag{26}$$

Proof: The proof follows from the ansatz 15. See A.

The value function above solves a general discrete-time stochastic optimal control problem where an agent α tracks an exogenous signal $\boldsymbol{v}_k^{\alpha}$. The problem is intractable in general as it requires the computation of the expectation of the offset, $\mathbb{E}_k[\boldsymbol{s}_{k+1}^{\alpha}]$. However, for the optimal strategy with the value function V_k^{α} to give a Nash equilibrium for the overall game, at each time step k the chosen strategy must generate the local field term $\boldsymbol{z}$ to be tracked, i.e. $\boldsymbol{u}_k^o$ must generate a trajectory satisfying $\boldsymbol{z}_k^{\alpha} = \boldsymbol{M}\boldsymbol{x}_k^{\alpha} = \boldsymbol{v}_k^{\alpha}$ for all α. This is known as the consistency condition for the Nash equilibrium in the limit game [8]. In the full state feedback case, the consistency condition allows the expectation of the offset $\boldsymbol{s}_k$ to be explicitly calculated.

B. Nash Equilibrium Consistency Condition with Full State Information

By definition, for each k, the local field $\boldsymbol{z}_k$ is given by $\boldsymbol{z}_k = \boldsymbol{M}\boldsymbol{x}_k$. As $\boldsymbol{x}_k$ is square-integrable for each k when generated by the optimal strategy $\boldsymbol{u}_k$ and $\boldsymbol{M}$ is an $L_2[0,1]$ to $L_2[0,1]$ operator, $\boldsymbol{z}_k \in L_2[0,1]$. For the game to yield a Nash equilibrium, it is necessary for all agents to apply their respective control $\boldsymbol{u}_k^{\alpha}$ generating the local field process $\boldsymbol{z}_k^{\alpha}$. To denote the function over the whole index set the superscript α is omitted.

Lemma 4.2: Let the signal to be tracked be given by $\boldsymbol{z}_k = \boldsymbol{M}\boldsymbol{x}_k$ for time k. Let Γ_k and Ψ_k be $L_2([0,1])$ operators which are defined by the backwards recursion equations

$$\begin{aligned} \Psi_k =& - S\mathbb{I} + F_k R(G_k + H_k\Psi_{k+1}\Gamma_k) \\ &+ \frac{1}{2}(A - BF_k)^* P_{k+1}\big[(D - BH_k) \\ &- B\Psi_{k+1}H_k\Gamma_k\big] \\ &+ (A - BF_k)^*\Psi_{k+1}\Gamma_k, \end{aligned} \tag{27}$$

$$\begin{aligned} \Gamma_k =& (\mathbb{I} + BH_k\boldsymbol{M}\Psi_{k+1})^{-1}[(A - BF_k)\mathbb{I} \\ &+ (D - BG_k)\boldsymbol{M}] \end{aligned} \tag{28}$$

with the terminal conditions

$$\Psi_T = - S\mathbb{I}, \tag{29}$$

$$\begin{aligned} \Gamma_{T-1} =& (\mathbb{I} - SBH_{T-1}\boldsymbol{M})^{-1}[(A - BF_{T-1})\mathbb{I} \\ &+ (D - BG_{T-1})\boldsymbol{M}]. \end{aligned} \tag{30}$$

Assume that for all $k = \{0,...,T-1\}$, the inverse $(\mathbb{I} + BH_k\boldsymbol{M}\Psi_{k+1})^{-1}$ exists. Then,

$$\mathbb{E}_k[\boldsymbol{z}_{k+1}] = \Gamma_k\boldsymbol{z}_k, \tag{31}$$

$$\boldsymbol{s}_k = \Psi_k\boldsymbol{z}_k, \tag{32}$$

and the trajectory generated by

$$\boldsymbol{u}_k = -F_k \boldsymbol{x}_k - (G_k \mathbb{I} + H_k \Psi_{k+1} \Gamma_k) \boldsymbol{z}_k \tag{33}$$

gives the optimal tracking trajectory for each α.

Proof: See B.

Combining Lemma A and B yields the Nash equilibrium of the game.

Theorem 4.3: Given the limit graphon tracking game of the type (11) for the family of systems (12), where each agent α indexed by $[0, 1]$ has the information pattern $\mathcal{F}_k^\alpha = \{\boldsymbol{x}_k, \boldsymbol{z}_k\}$, the control strategy given in equations (27), (28), and (33) yields a Nash equilibrium.

V. Numerical Simulation

The behavior of the full information Nash equilibrium system is demonstrated by simulating a large network averaging game of $N = 500$ agents with a scalar state, where each agent is given a position on the unit interval uniformly. Set $A = 0.5$, $B = 1$, and $D = 1$. The connection strength of the network is approximated by the graphon $\boldsymbol{M}(\alpha, \beta) = \cos(\pi(\alpha - \beta))$, evaluated on the points of a uniform grid of partition length 0.002. Set the initial state $\boldsymbol{x}_0^\alpha = 1$ for all agents.

Set the state cost $S = 2$ and the control cost $R = 1$, and the initial condition for all agents to be equal to one. To investigate the impact of the Q-noise intensity, a system generating a Nash equilibrium is tested where the covariance of the disturbance is given by $\mathbf{Q}_1(\alpha, \beta) = 1 - \max(\alpha, \beta)$ in Fig. 3, $\mathbf{Q}_2(\alpha, \beta) = (1 - \max(\alpha, \beta))/10$ in Fig. 4, and $\mathbf{Q}_3(\alpha, \beta) = (1 - \max(\alpha, \beta))/100$ in Fig. 5.

By decreasing the noise intensity, the general structure of the Nash equilibrium can be seen. Without the field tracking penalty, each agent would independently drive its state $\boldsymbol{x}_k^\alpha$ to zero. The field tracking behavior modifies the rate at which an agent drives their local state to zero. This behavior is evident in Fig. 4-II and Fig. 5-II.

Once the local states of all agents are near the origin, the dynamics are dominated by the noise. This is apparent in Fig. 3-III, as the random perturbations to the state of each agent overshadow the agents' control. The decreasing maximum and minimum values of Fig. 4-III and Fig. 5-III confirm this as the system is less impacted by the noise, and hence the agents' local state $\boldsymbol{x}_k^\alpha$ is closer to their field term $\boldsymbol{z}_k^\alpha$ at the terminal time.

VI. Future work

There are some immediate directions for future research. First, the work should be extended to limit graphs embedded in metric spaces using the embedded graph limit theory developed in Caines [12]. This theory generalizes the concept used implicitly in the numerical simulations above, where each node in the graph is located uniformly at a point on the unit interval. Embedded graph limit theory is a method for describing graph limits that exist in geometric spaces more general than the unit interval, for instance those where each node is located in $\mathbb{R}^2$ or $\mathbb{R}^3$.

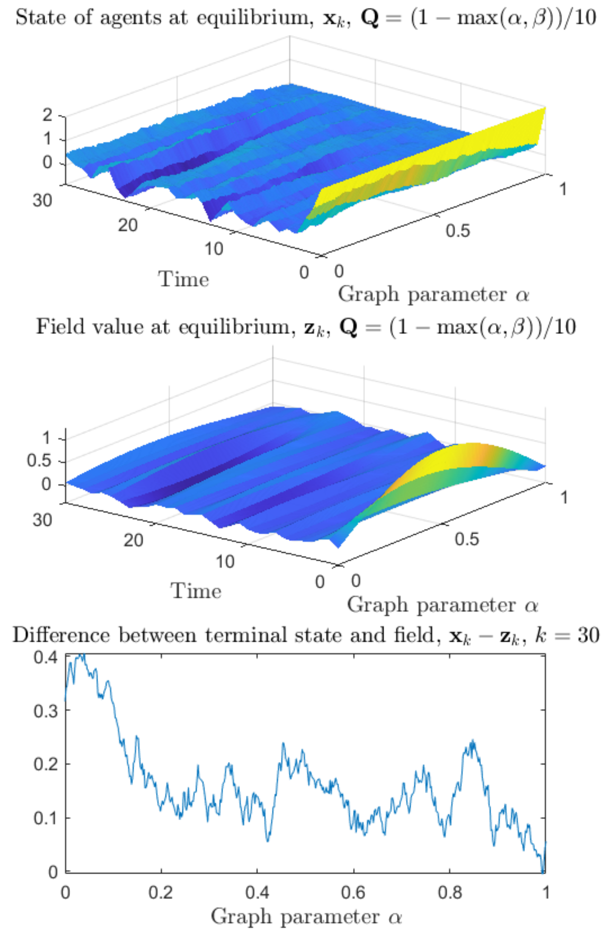

Fig. 3. I: the state trajectory of a single sample path of the graphon field tracking game with $\mathbf{Q} = 1 - \max(\alpha, \beta)$, compensating for the field term using Lemma 4.1. II: the graphon field being tracked at Nash equilibrium. III: the difference between the terminal state $\boldsymbol{x}_k$ and the terminal field $\boldsymbol{z}_k$, $k = 30$. The error present in this image is primarily a result of the high noise intensity in the system.

The impact of different forms of the Q-noise covariance on the resulting tracking game should also be further developed.

This article considered Nash equilibria with full state information. This will be expanded to other information sets, such as those where each agent has only local information and, hence, estimation of the status of the overall graphon field may be of value.

References

[1] M. Huang, P. E. Caines, and R. P. Malhame, "The NCE (Mean Field) Principle With Locality Dependent Cost Interactions," *IEEE Transactions on Automatic Control*, vol. 55, no. 12, pp. 2799–2805, Dec. 2010, conference Name: IEEE Transactions on Automatic Control.

[2] A. Bensoussan, J. Frehse, and P. Yam, *Mean Field Games and Mean Field Type Control Theory*, ser. SpringerBriefs in Mathematics. New York, NY: Springer, 2013. [Online]. Available: http://link.springer.com/10.1007/978-1-4614-8508-7

[3] R. Carmona and F. Delarue, *Probabilistic Theory of Mean Field Games with Applications II*, 2018. [Online].

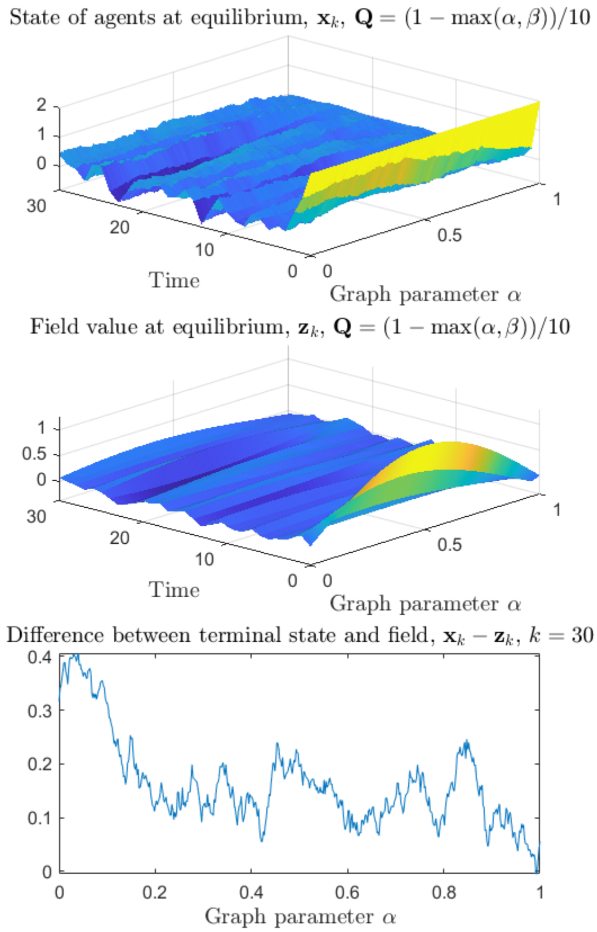

Fig. 4. I: the state trajectory of a single sample path of the graphon field tracking game with $\mathbf{Q} = (1 - \max(\alpha, \beta))/10$, compensating for the field term using Lemma 4.1. II: the graphon field being tracked at the corresponding Nash equilibrium. III: the difference between the terminal state $\boldsymbol{x}_k$ and the terminal field $\boldsymbol{z}_k$, $k = 30$. The error magnitude is lower than in Fig. 3-III, indicating that the agents more effectively track the field when there is lower noise intensity.

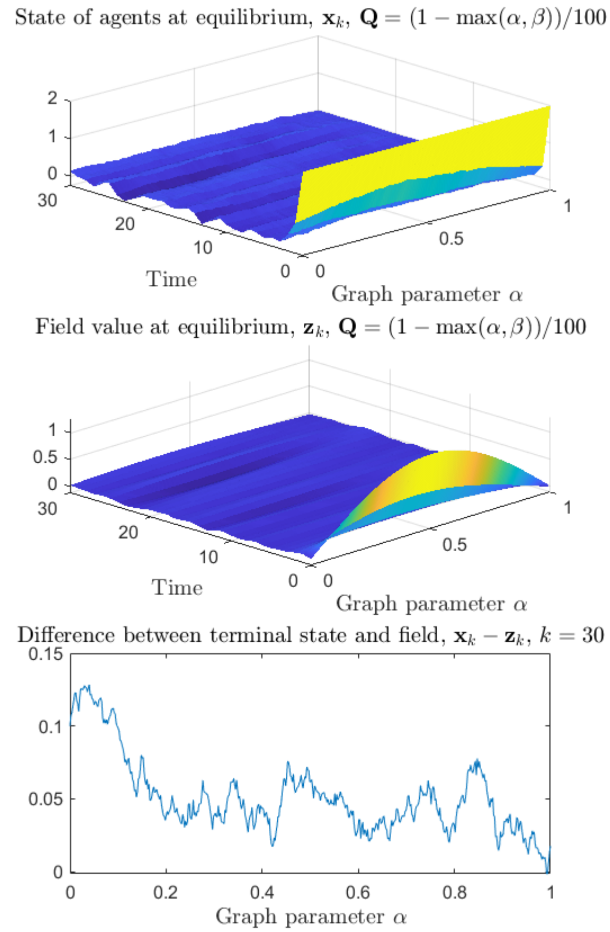

Fig. 5. I: the state trajectory of a single sample path of the graphon field tracking game with $\mathbf{Q} = (1 - \max(\alpha, \beta))/100$, compensating for the field term using Lemma 4.1. II: The graphon field being tracked at Nash equilibrium. III: the difference between the terminal state $\boldsymbol{x}_k$ and the terminal field $\boldsymbol{z}_k$, $k = 30$. As expected, when the noise intensity is low, the state of each agent tends to zero.

Available: https://link-springer-com.proxy3.library.mcgill.ca/book/10.1007/978-3-319-56436-4

[4] P. E. Caines and M. Huang, "Graphon Mean Field Games and the GMFG Equations," in *2018 IEEE Conference on Decision and Control (CDC)*, Dec. 2018, pp. 4129–4134, iSSN: 2576-2370.

[5] ——, "Graphon Mean Field Games and the GMFG Equations: Epsilon-Nash Equilibria," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, Dec. 2019, pp. 286–292, iSSN: 2576-2370.

[6] ——, "Graphon Mean Field Games and Their Equations," *SIAM Journal on Control and Optimization*, vol. 59, no. 6, pp. 4373–4399, Jan. 2021, publisher: Society for Industrial and Applied Mathematics. [Online]. Available: http://epubs.siam.org/doi/10.1137/20M136373X

[7] L. Lovász, *Large Networks and Graph Limits*, 2012. [Online]. Available: https://bookstore.ams.org/coll-60/

[8] S. Gao, R. F. Tchuendom, and P. E. Caines, "Linear quadratic graphon field games," *Communications in Information and Systems*, vol. 21, no. 3, pp. 341–369, Jun. 2021, publisher: International Press of Boston. [Online]. Available: https://content.intlpress.com/journal/CIS/article/545

[9] A. Dunyak and P. E. Caines, "Linear Stochastic Graphon Systems with Q-Space Noise," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, Dec. 2022, pp. 3926–3932, iSSN: 2576-2370.

[10] A. Kukush, *Gaussian Measures in Hilbert Space*. Wiley, 2019. [Online]. Available: http://onlinelibrary.wiley.com/doi/epub/10.1002/9781119476825

[11] S. Gao and P. E. Caines, "Graphon Control of Large-Scale Networks of Linear Systems," *IEEE Transactions on Automatic Control*, vol. 65, no. 10, pp. 4090–4105, Oct. 2020, conference Name: IEEE Transactions on Automatic Control.

[12] P. E. Caines, "Embedded Vertexon-Graphons and Embedded GMFG Systems," in *IEEE Conference on Decision and Control*, Dec. 2022.

APPENDIX

A. Proof of Lemma 4.1

The dynamic programming principle is applied to find the optimal control. From the terminal condition

$$V_T^{\alpha}(\boldsymbol{x}_k) = ||\boldsymbol{x}_T^{\alpha} - \boldsymbol{v}_T^{\alpha}||_S^2. \tag{34}$$

Then, $P_T = S$, $\boldsymbol{s}_T^{\alpha} = -S\boldsymbol{v}_T^{\alpha}$, and $m_T^{\alpha} = ||\boldsymbol{v}_T^{\alpha}||_S^2$.

By the dynamic programming assumption,

$$V_k^\alpha(\boldsymbol{x}_k) = \min_u \mathbb{E}_k\left[||\boldsymbol{x}_k^\alpha - \boldsymbol{v}_k^\alpha||_S^2 + ||u||_R^2 + V_{k+1}^\alpha(\boldsymbol{x}_{k+1})\right] \tag{35}$$

$$= \min_u ||\boldsymbol{x}_k^\alpha - \boldsymbol{v}_k^\alpha||_S^2 + ||u||_R^2 + \mathbb{E}_k[V_{k+1}^\alpha(\boldsymbol{x}_{k+1})] \tag{36}$$

$$\begin{aligned}= \min_u &||\boldsymbol{x}_k^\alpha - \boldsymbol{v}_k^\alpha||_S^2 + ||u||_R^2 \\ &+ \mathbb{E}_k[(\boldsymbol{x}_{k+1}^\alpha)^* P_{k+1}(\boldsymbol{x}_{k+1}^\alpha) + 2(\boldsymbol{x}_{k+1}^\alpha)^*\boldsymbol{s}_{k+1}^\alpha \\ &\quad + m_{k+1}^\alpha]\end{aligned} \tag{37}$$

$$\begin{aligned}= \min_u &||\boldsymbol{x}_k^\alpha - \boldsymbol{v}_k^\alpha||_S^2 + ||u||_R^2 \\ &+ \mathbb{E}_k[(A\boldsymbol{x}_k^\alpha + B\boldsymbol{u}_k^\alpha + D\boldsymbol{v}_k^\alpha + \boldsymbol{g}_k^\alpha)^* P_{k+1} \\ &\cdot (A\boldsymbol{x}_k^\alpha + B\boldsymbol{u}_k^\alpha + D\boldsymbol{v}_k^\alpha + \boldsymbol{g}_k^\alpha)] \\ &+ 2\mathbb{E}_k[A\boldsymbol{x}_k^\alpha + B\boldsymbol{u}_k^\alpha + D\boldsymbol{v}_k^\alpha + \boldsymbol{g}_k^\alpha]\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha] \\ &+ \mathbb{E}_k[m_{k+1}^\alpha]\end{aligned} \tag{38}$$

$$\begin{aligned}= \min_u &||\boldsymbol{x}_k^\alpha - \boldsymbol{v}_k^\alpha||_S^2 + ||u||_R^2 \\ &+ (A\boldsymbol{x}_k^\alpha + B\boldsymbol{u}_k^\alpha + D\boldsymbol{v}_k^\alpha)^* P_{k+1} \\ &\cdot (A\boldsymbol{x}_k^\alpha + B\boldsymbol{u}_k^\alpha + D\boldsymbol{v}_k^\alpha) + \mathbf{Q}(\alpha,\alpha) \\ &+ 2(A\boldsymbol{x}_k^\alpha + B\boldsymbol{u}_k^\alpha + D\boldsymbol{v}_k^\alpha)\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha] + \mathbb{E}_k[m_{k+1}^\alpha].\end{aligned} \tag{39}$$

Note that the right-hand expression of (39) is differentiable and convex in u, and hence the optimal control is

$$\begin{aligned}u_k^{o,\alpha} =& -(R + B^* P_{k+1} B)^{-1} B^*[P_{k+1}(A\boldsymbol{x}_k^\alpha + D\boldsymbol{v}_k^\alpha) \\ &+ \mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]]\end{aligned} \tag{40}$$

$$=: -F_k\boldsymbol{x}_k^\alpha - G_k\boldsymbol{v}_k^\alpha - H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}^\alpha]. \tag{41}$$

Applying the optimal control to the value function and rearranging terms gives equations (16–21) as required. □

B. Proof of Lemma 4.2

First, recall that by definition $\boldsymbol{z}_k = \boldsymbol{M}\boldsymbol{x}_k$, and hence when applying the optimal control at time $k = T-1$,

$$\mathbb{E}_{T-1}[\boldsymbol{z}_T] = \mathbb{E}_{T-1}[\boldsymbol{M}\boldsymbol{x}_T] \tag{42}$$

$$= \mathbb{E}_{T-1}[\boldsymbol{M}(A\boldsymbol{x}_{T-1} + B\boldsymbol{u}_{T-1} + D\boldsymbol{z}_{T-1} + \boldsymbol{g}_{T-1})]] \tag{43}$$

$$\begin{aligned}=& \boldsymbol{M}[A\boldsymbol{x}_{T-1} + B(-F_{T-1}\boldsymbol{x}_{T-1} \\ &- G_{T-1}\boldsymbol{z}_{T-1} - H_{T-1}\mathbb{E}_{T-1}[\boldsymbol{s}_T]) + D\boldsymbol{z}_{T-1}]\end{aligned} \tag{44}$$

$$\begin{aligned}=& \boldsymbol{M}[(A - BF_{T-1})\boldsymbol{x}_{T-1} + (D - BG_{T-1})\boldsymbol{z}_{T-1} \\ &- BH_{T-1}\mathbb{E}_{T-1}[\boldsymbol{s}_T]]\end{aligned} \tag{45}$$

$$\begin{aligned}=& (A - BF_{T-1})\boldsymbol{M}\boldsymbol{x}_{T-1} \\ &+ (D - BG_{T-1})\boldsymbol{M}\boldsymbol{z}_{T-1} - BH_{T-1}\boldsymbol{M}\mathbb{E}_{T-1}[\boldsymbol{s}_T]\end{aligned} \tag{46}$$

$$\begin{aligned}=& (A - BF_{T-1})\boldsymbol{z}_{T-1} + (D - BG_{T-1})\boldsymbol{M}\boldsymbol{z}_{T-1} \\ &- BH_{T-1}\boldsymbol{M}\mathbb{E}_{T-1}[\boldsymbol{s}_T].\end{aligned} \tag{47}$$

Then, applying the terminal condition $\boldsymbol{s}_T = -S\boldsymbol{z}_T$,

$$\begin{aligned}\mathbb{E}_{T-1}[\boldsymbol{z}_T] =& (A - BF_{T-1})\boldsymbol{z}_{T-1} + (D - BG_{T-1})\boldsymbol{M}\boldsymbol{z}_{T-1} \\ &+ BH_{T-1}\boldsymbol{M}\mathbb{E}_{T-1}[S\boldsymbol{z}_T]\end{aligned} \tag{48}$$

$$\begin{aligned}\mathbb{E}_{T-1}[\boldsymbol{z}_T] - BH_{T-1}\boldsymbol{M}\mathbb{E}_{T-1}[S\boldsymbol{z}_T] =& (A - BF_{T-1})\boldsymbol{z}_{T-1} \\ &+ (D - BG_{T-1})\boldsymbol{M}\boldsymbol{z}_{T-1}.\end{aligned} \tag{49}$$

Hence,

$$\mathbb{E}_{T-1}[\boldsymbol{z}_T] = (\mathbb{I} - SBH_{T-1}\boldsymbol{M})^{-1}[(A - BF_{T-1})\mathbb{I} \tag{50}$$

$$+ (D - BG_{T-1})\boldsymbol{M}]\boldsymbol{z}_{T-1} \tag{51}$$

$$=: \Gamma_{T-1}\boldsymbol{z}_{T-1}. \tag{52}$$

Observing this, make the following inductive hypothesis:

$$\mathbb{E}_k[\boldsymbol{z}_{k+1}] = \Gamma_k\boldsymbol{z}_k, \tag{53}$$

$$\boldsymbol{s}_k = \Psi_k\boldsymbol{z}_k, \tag{54}$$

where Ψ_k and Γ_k are $L_2([0,1])$ operators for each $k \in \{0, ..., T\}$. Applying the inductive hypotheses to the expectation of $\boldsymbol{z}_{k+1}$,

$$\begin{aligned}\mathbb{E}_k[\boldsymbol{z}_{k+1}] =& [(A - BF_k)\mathbb{I} + (D - BG_k)\boldsymbol{M}]\boldsymbol{z}_k \\ &- BH_k\boldsymbol{M}\mathbb{E}_k[\boldsymbol{s}_{k+1}]\end{aligned} \tag{55}$$

$$\begin{aligned}=& [(A - BF_k)\mathbb{I} + (D - BG_k)\boldsymbol{M}]\boldsymbol{z}_k \\ &- BH_k\boldsymbol{M}\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}]\end{aligned} \tag{56}$$

$$\begin{aligned}=& (\mathbb{I} + BH_k\boldsymbol{M}\Psi_{k+1})^{-1} \\ &\cdot [(A - BF_k)\mathbb{I} + (D - BG_k)\boldsymbol{M}]\boldsymbol{z}_k\end{aligned} \tag{57}$$

$$=: \Gamma_k\boldsymbol{z}_k, \tag{58}$$

which shows equation (53). Applying the inductive hypotheses to the recursion for $\boldsymbol{s}_k$,

$$\begin{aligned}\boldsymbol{s}_k =& -S\boldsymbol{z}_k + F_k^* R(G_k\boldsymbol{z}_k + H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}]) \\ &+ \frac{1}{2}(A - BF_k)^* P_{k+1} \\ &\cdot \left[(D - BH_k)\boldsymbol{z}_k - BH_k\mathbb{E}_k[\boldsymbol{s}_{k+1}]\right] \\ &+ (A - BF_k)^*\mathbb{E}_k[\boldsymbol{s}_{k+1}]\end{aligned} \tag{59}$$

$$\begin{aligned}=& -S\boldsymbol{z}_k + F_k^* R(G_k\boldsymbol{z}_k + H_k\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}]) \\ &+ \frac{1}{2}(A - BF_k)^* P_{k+1}\left[(D - BH_k)\boldsymbol{z}_k \right.\\ &\left.- BH_k\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}]\right] \\ &+ (A - BF_k)^*\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}]\end{aligned} \tag{60}$$

$$\begin{aligned}=& -S\boldsymbol{z}_k + F_k R(G_k\boldsymbol{z}_k + H_k\Psi_{k+1}\mathbb{E}_k[\boldsymbol{z}_{k+1}]) \\ &+ \frac{1}{2}(A - BF_k)^* P_{k+1}\left[(D - BH_k)\boldsymbol{z}_k \right.\\ &\left.- B\Psi_{k+1}H_k\mathbb{E}_k[\boldsymbol{z}_{k+1}]\right] \\ &+ (A - BF_k)^*\Psi_{k+1}\mathbb{E}_k[\boldsymbol{z}_{k+1}]\end{aligned} \tag{61}$$

$$\begin{aligned}=& -S\boldsymbol{z}_k + F_k R(G_k\boldsymbol{z}_k + H_k\Psi_{k+1}\Gamma_k\boldsymbol{z}_k) \\ &+ \frac{1}{2}(A - BF_k)^* P_{k+1}\left[(D - BH_k)\boldsymbol{z}_k \right.\\ &\left.- B\Psi_{k+1}H_k\Gamma_k\boldsymbol{z}_k\right] \\ &+ (A - BF_k)^*\Psi_{k+1}\Gamma_k\boldsymbol{z}_k\end{aligned} \tag{62}$$

$$=: \Psi_k\boldsymbol{z}_k. \tag{63}$$

Then, the optimal control $\boldsymbol{u}_k^o$ is given in

$$\boldsymbol{u}_k^o = -F_k\boldsymbol{x}_k - G_k\boldsymbol{z}_k - H_k\mathbb{E}_k[\boldsymbol{s}_{k+1}] \tag{64}$$

$$= -F_k\boldsymbol{x}_k - G_k\boldsymbol{z}_k - H_k\mathbb{E}_k[\Psi_{k+1}\boldsymbol{z}_{k+1}] \tag{65}$$

$$= -F_k\boldsymbol{x}_k - (G_k\mathbb{I} + H_k\Psi_{k+1}\Gamma_k)\boldsymbol{z}_k. \tag{66}$$

□